Probing the Role of Gag in Regulation of Reverse Transcription

Andrew Goodrich

Departments of Biochemistry and Chemistry

The Ohio State University

Columbus, OH 43210

May 2010

Abstract

The HIV-1 nucleocapsid (NC) domain of Gag has both specific and more general nucleic acid (NA) binding properties. NC acts specifically to recognize and package unspliced viral RNA and as a general chaperone to facilitate reverse transcription by annealing/ aggregating NAs via its highly positive character and by destabilizing NA secondary structure via its two CCHC zinc finger (ZF) motifs. Interestingly, it has recently been reported that ZF deletion or mutation to CCCC results in production of virions containing DNA instead of RNA, rendering them noninfectious. Thus, an additional role of NC is to prevent premature reverse transcription from occurring before or during assembly. Here, we probe the *in vitro* NA binding and chaperone properties of Gag variants containing the same ZF mutations or deletions tested in the cell-based assays. Fluorescence anisotropy equilibrium binding measurements reveal that mutation or deletion of both ZFs results in a modest reduction in binding to the ψ SL3 stem-loop relative to WT Gag, whereas binding to nonspecific single-stranded NAs is largely unaffected. Similarly, Gag's ability to aggregate NAs or facilitate $tRNA^{Lys,3}$ annealing to the primer-binding site, two functions of Gag during viral assembly, was only moderately affected upon ZF mutation or deletion. A time-resolved fluorescence resonance energy transfer assay was used to monitor hairpin stem opening of a DNA hairpin construct. Surprisingly, single CCCC and ZF deletion variants appeared to be more effective at opening the TAR hairpin than WT Gag, and Gag variants in which both zinc fingers were mutated or deleted showed even greater duplex destabilization capability than the single ZF variants. Previous studies with the freestanding NC domain of Gag indicated that NC's duplex destabilization activity depended on the ZF structures. Our new results suggest that in the context of Gag, disruption of the ZF leads to an increased ability to disrupt nucleic acid secondary structure and we explore possible mechanisms by which this increased capability may cause premature reverse transcription.

Introduction

Human immunodeficiency virus (HIV), the causative agent of acquired immune deficiency syndrome (AIDS), has been among the greatest global health concerns in the past three decades. A retrovirus, HIV carries its genetic material as two copies of single-stranded RNA (ssRNA) which must be reverse transcribed by the viral enzyme reverse transcriptase (RT) into double-stranded DNA (dsDNA) after infection of a new cell (Fig. 1)[4]. The dsDNA genome is then integrated into the host cell's DNA by the viral integrase (IN), thereby establishing a permanent infection. The integrated proviral DNA is transcribed and translated by the host cell machinery. The newly transcribed ssRNA genomes then assemble with the viral

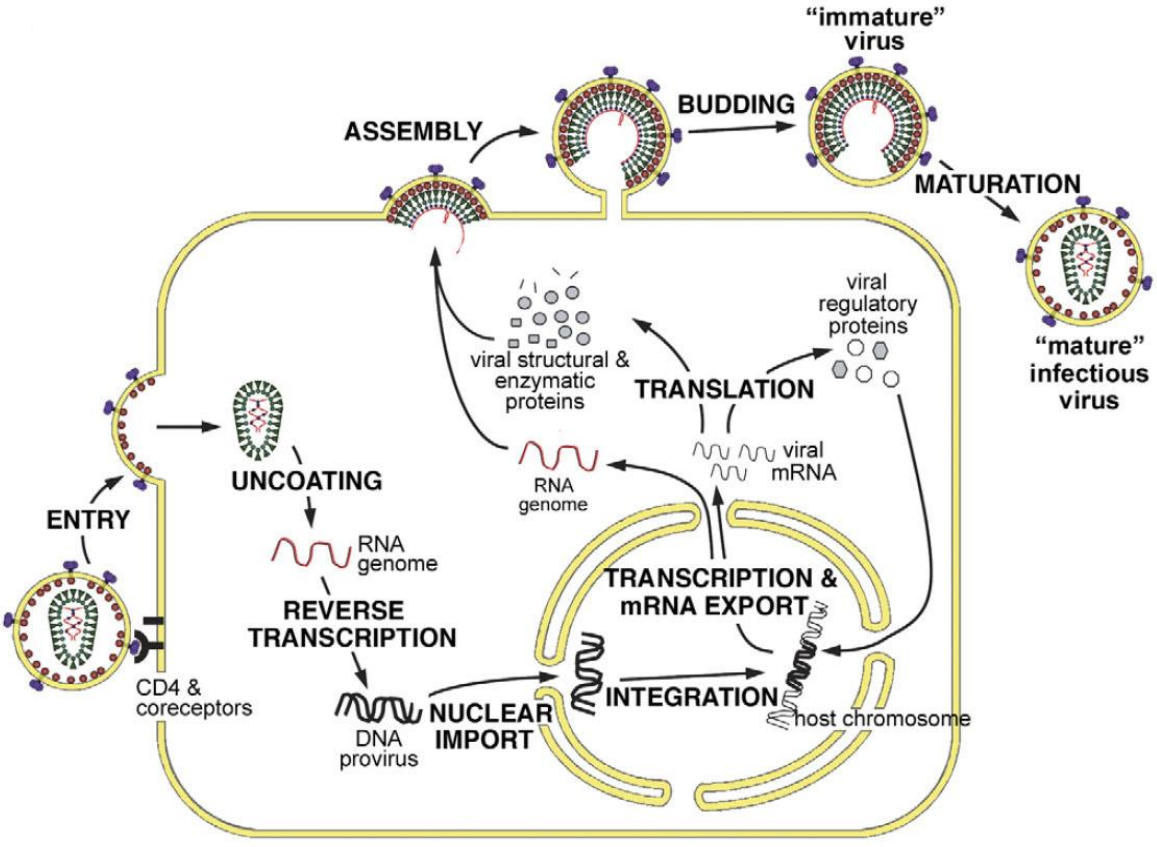

FIGURE 1. HIV life cycle. Figure adapted from [2].

proteins at the plasma membrane and bud from the cell to create new virions. During maturation, the viral protease (PR) cleaves the precursor proteins into their domains, allowing the mature virus particles to propagate the infection.

Due to the structural complexity of the ssRNA viral genome, HIV relies on nucleic acid (NA) chaperones to facilitate many steps in its life cycle [5-6]. NA chaperones are proteins that help prevent RNAs from misfolding and facilitate formation of the thermodynamically most favorable folded state [7]. The critical aspects of chaperone activity are the abilities to aggregate NAs and destabilize NA secondary structure; fast on/off binding kinetics play an important role, too [8]. HIV-1 nucleocapsid (NC) is a small (55-residue), basic protein which contains two highly conserved CCHC-type zinc fingers separated by a short linker domain, a highly basic N-terminal tail, and a short C-terminal tail. The two zinc fingers are critical to NC's ability to destabilize NA secondary structure, with the N-terminal finger being more important than the C-terminal finger, while aggregation activity is attributed to the highly cationic, unstructured N-terminal region. NC is the primary NA chaperone for most of the viral life cycle. NC begins as a part of the Gag polyprotein and is released as the free-standing NC protein following proteolysis

FIGURE 2. Schematic representation of Gag. Its domains from N- to C-terminus are matrix (MA), capsid (CA), spacer peptide 1 (SP1), nucleocapsid (NC), spacer peptide 2 (SP2), and p6. Arrows indicate sites of cleavage by protease (PR) during maturation. Adapted from [2].

of the polyprotein precursor Gag. NC or the NC domain of Gag have been shown to be critical to many steps of the viral life cycle, including assembly, reverse transcription, and integration [6,9-10].

Gag, which consists of the matrix (MA), capsid (CA), SP1, NC, SP2, and P6 domains, is the main structural protein in HIV and is necessary and sufficient for virus particle formation *in vitro* [2]. Gag is responsible for packaging two copies of the ssRNA genome, annealing the $tRNA^{Lys3}$ to the primer binding site (PBS) near the 5' end of the genome [9], and promoting dimerization of the genome [9].

More recently, reports have suggested that Gag may also play an important role in regulating the timing of reverse transcription [10-13]. Gag variants containing mutations to critical zinc-binding residues within the zinc fingers have long been recognized as being replication defective [14]. It was only recently realized that these mutations allow for reverse transcription to occur prematurely in virus-producing cells prior to packaging, leading to the packaging of reverse transcription intermediates being packaged into newly forming virions [10-11]. Early reverse transcription products were detected in almost every single virion, representing a 1,000-fold increase in intravirion DNA [10]. The dsDNA genomes that result from this premature reverse transcription are inherently less stable than those resulting from wild-type virions and are unable to integrate their genomes into newly infected cells [10-11]. Thus, it would seem that properly timed reverse transcription is a necessary prerequisite for HIV infection and that the architecture of the zinc fingers within the NC domain of Gag play a critical role in determining when reverse transcription begins. However, the mechanism by which Gag exerts this temporal control has not been established.

In this study, we purified each of the Gag variants resulting in the premature reverse transcription phenotype and examined their NA chaperone activity *in vitro* using a series of biochemical and biophysical assays. Insights gained from these experiments not only shed light

on Gag's role in regulating the timing of reverse transcription but also represent the first mutational analysis of NC within the context of Gag.

Materials and Methods

Protein and Nucleic Acid Preparation

In vitro transcribed human $tRNA^{Lys3}$ and ShortPBS RNA[15], a 105-nt RNA taken from the 5' UTR of the HIV-1 genome containing the primer binding site, were prepared as previously described.

Mutant Gag sequences were constructed using the Site-Directed Mutagenesis Kit from Stratagene on pET3xc coding for GagΔP6 (A gift from Dr. Alan Rein). Primers used to generate mutant plasmids are listed in Table 1 along with plasmid names.

Table 1. Primer Design for Mutagenesis of pET3xc

pET3xcΔZF1	
ΔZF1 top	
CCAAAGAAAGATGGTTAAGAGGGCCCCTAGG	
ΔZF1 bottom	
CCTAGGGGCCCTCTTAACCATCTTTCTTTGG	
pET3xcΔZF2	
ΔZF2 top	
CCCCTAGGAAAAAGGGCACTGAGAGACAGGC	
ΔZF2 bottom	
GCCTGTCTCTCAGTGCCCTTTTTCCTAGGGG	
pET3xcH400C	
H400C top	
GGCAAAGAAGGGTGCACAGCCAGAAATTGCAGGG	
H400C bottom	
CCCTGCAATTTCTGGCTGTGCACCCTTCTTTGCC	
pET3xcH421C	
H421C top	
GTGGAAAGGAAGGATGCCAAATGAAAGATTGTACTG	
H421C bottom	
CAGTACAATCTTTCATTTGGCATCCTTCCTTTCCAC	
Plasmid coding for ΔZF1+2 referred to as pET3xcΔZF1+2	
Plasmid coding for H400,421C referred to as pET3xcH400,421C	

pET3xcH400,421C was generated by performing site-directed mutagenesis upon pET3xcH400C using the H421C primers and pET3xcΔZF1+2 was generated by performing site-directed mutagenesis upon pET3xcΔZF2 using the ΔZF1 primers.

Proteins were overexpressed and purified from e. coli as previously described [16]. Briefly, protein was overexpressed in BL21 (DE3) pLysS cells. 200 ml starter cultures containing 100 μg/ml ampicillin and 34 μg/ml chloramphenicol were incubated overnight at 37

°C with shaking. After ~16 hours growth, the starter culture was divided evenly amongst 4 L LB + antibiotics and incubated at 37 °C with shaking until OD_{600} measured 0.6. Isopropyl-(β-D)thiogalactopyranoside (IPTG) was added to a final concentration of 100 μg/ml, and incubation with shaking continued for an additional 4-5 hours. Cells were harvested by centrifuging at 3,500 x g and pellets stored at -80 °C.

Gag ΔP6 and mutant proteins were purified as follows [16]. All steps were performed at 4 °C unless otherwise indicated. Pellets obtained from induction were resuspended in lysis buffer (20 mM Tris-HCl (pH 7.4), 1 mM PMSF, 10 mM β-mercaptoethanol, 1 μM $ZnCl_2$, 0.75 M NaCl, 10% glycerol w/v, 0.05% Tween-20 w/v). Resuspended pellets were then sonicated on ice until a homogenous lysate was achieved. The resulting solution was centrifuged at 12000 x g for 15 minutes. 0.5 volumes saturated ammonium sulfate were then added at a rate of ~0.75 ml/minute to the cleared lysate on ice with stirring, after which the resulting solution was allowed to precipitate for 30 minutes and centrifuged at 12,000 x g for 15 minutes. The pellet was resuspended, and insoluble material was removed. The solution was then added to phosphocellulose resin prepared as specified in [16]. The resin was washed, and Gag was eluted with a stepwise NaCl gradient from 0.1 M NaCl to 1 M NaCl. Gag eluted at 1 M NaCl. Ammonium sulfate was then added to eluents to a final concentration of 60% saturation and precipitated on ice overnight. Precipitated protein was pelleted by centrifuging at 12,000 x g for 15 minutes and redissolved in a minimal amount of storage buffer. The resulting solution was centrifuged at 20,000 x g for 15 minutes to remove any material that failed to dissolve. The supernatant was then transferred to a dialysis cassette (3-12 ml capacity, 10,000 MWCO) and dialyzed overnight against storage buffer (20 mM Tris-HCl (pH 7.4), 0.5 M NaCl, 10 mM β-mercaptoethanol, 0.05% v/v tween-20, 10 % w/v glycerol, 1 μM $ZnCl_2$). Znless Gag was

prepared by purifying WT Gag ΔP6 as described and dialyzing against storage buffer containing 10 mM EDTA before dialyzing twice against storage buffer without $ZnCl_2$. Concentration was estimated by measuring the absorbance at 280 nm using ε_{280} of $6.3x10^4$ $M^{-1}*cm^{-1}$ for Gag ΔP6, Znless Gag, H400C, H421C, H400,421C, and ΔZF1, $5.8x10^4$ $M^{-1}*cm^{-1}$ for ΔZF2, and $5.7x10^4$ for ΔZF1+2, as determined using the Scripps Protein Calculator. Protein was then aliquoted and stored at -80 °C.

Annealing Assays

ShortPBS RNA and $tRNA^{Lys3}$ were folded prior to use as follows [15]. ShortPBS RNA was resuspended in 50 mM HEPES, heated at 85 °C for 2.5 mins., 50 °C for 8 mins., $MgCl_2$ added to a final concentration of 10 mM, heated at 37 °C for 10 mins., and placed on ice for at least 30 mins. before use. $tRNA^{Lys3}$ was resuspended in 50 mM HEPES, heated at 80 °C for 2 mins., 60 °C for 2 mins., $MgCl_2$ added to a final concentration of 10 mM, and placed on ice for at least 30 mins. before use. Annealing reaction mixtures contained 50 mM HEPES (pH 7.4), 5 mM DTT, 1 mM $MgCl_2$, 150 mM NaCl, 10 nM $tRNA^{Lys3}$, and 25 nM Short PBS RNA. Reaction mixtures were incubated at 37 °C for 10 minutes prior to the addition of chaperone protein. For single time point assays reaction mixtures were incubated at 37 °C for 30 minutes. In time course annealing assays, 0.4 μM protein was used in the case of Gag ΔP6, H400C, H421C, ΔZF1, and ΔZF2, and 1 μM for ΔZF1+2, H400,421C, and Znless gag, and annealing was allowed to proceed at 37 °C for 30 minutes. In all annealing assays, reactions were quenched by adding SDS to a final concentration of 1%, phenol:chloroform extracted twice, mixed with loading dye (50% glycerol with dyes), and run on a polyacrylamide gel (5% stacking gel, 10% running gel). The gels were exposed to phosphor screens and visualized with a Typhoon Trio Variable Mode Imager (GE Healthcare). Single time point assays were fit to sigmoidal curves to determine

$K_{1/2,Anneal}$ values, and time course assays were fit to single exponential curves to determine reaction rate, k.

Sedimentation/Aggregation Assays

Reaction mixtures were identical to those for annealing assays. However, instead of quenching reactions, reactions were centrifuged at 10,000 RPM for 10 minutes to pellet aggregated NA-protein complexes. The supernatant, 2-4 µl, was analyzed using a scintillation counter. A plot of counts per minute versus protein concentration was fitted to a sigmoidal curve to determine $K_{1/2,Agg}$.

Fluorescence Anisotropy Assays

Fluorescence anisotropy was used to measure equilibrium binding of Gag ΔP6 and mutants to nucleic acids [17]. Reaction mixtures contained 10 nM fluorescently labeled NA in 150 mM NaCl, 20 mM HEPES (pH 7.4), 10 µM tris(2-carboxyethyl)phosphine, and 5 mM β-mercaptoethanol, and 1 mM $MgCl_2$. Alexafluor-488 labeled oligonucleotides were used in these studies: GA-22, 5'-Alexafluor-488-GAGAGAGAGAGAGAGAGAGAGA-3' (22nt), and SL3, 5'-GGACUAGCGGAGGCUAGUCCU-Alexafluor-488-3'(21nt), both purchased from Dharmacon. Measurements were obtained using a SpectraMax M5 plate reader (Molecular Devices, Sunnyvale, CA) according to established protocols. Binding affinities were determined by fitting the data to a 1:1 binding model with a correction for changes in fluorophore intensity due to protein binding.

Time-resolved FRET Assays

Time-resolved FRET assays were performed as previously described [18]. Briefly, a 64-nt DNA hairpin mimicking the TAR hairpin from the 5' end of the HIV-1 genome was labeled at the 5' end with Alexafluor 488 and at the 3' end with the fluorescence quencher 4-(4'dimethyl-aminophenylazo)benzoic acid was incubated with and without WT or mutant Gag for 30 mins. at room temperature in a buffer containing 20 mM HEPES, pH 7.5, 150 mM NaCl, 10 μM TCEP, 5 mM β-mercaptoethanol. After incubation, fluorescence lifetimes were measured, deconvoluted, and fit to a triple-exponential. From the fit, the relative populations of the closed, semi-open, and fully open states were determined. Relative free energy of the fully open state was calculated using ΔG=-RTln(% population of open state).

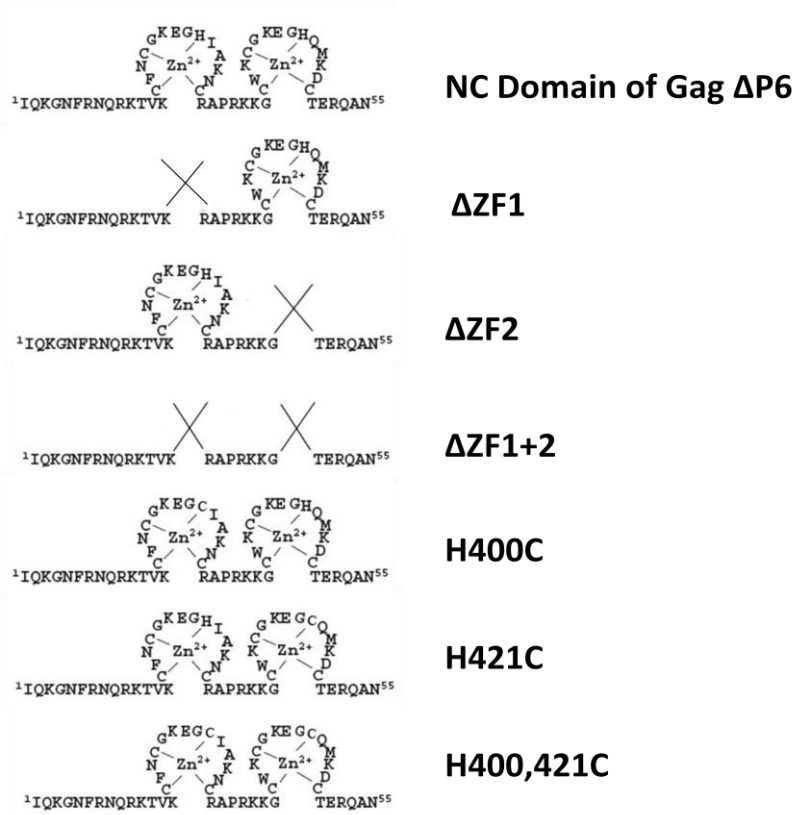

FIGURE 3. Sequence and zinc finger structure of the nucleocapsid domains of the WT HIV-1 Gag and mutant Gag proteins used in this study. Deletions eliminated residues from the first cysteine to the last cysteine within the CCHC zinc-binding motif. Point mutants resulted in mutation of zinc-binding histidine to cysteine, changing the zinc fingers from CCHC to CCCC. EDTA-treated Gag, also called Znless, has the same amino acid sequence as WT. Adapted from [12].

Results

Fig. 3 compares the amino acid sequence of the nucleocapsid domain of WT Gag ΔP6 and the Gag mutants used in this study. All mutants maintained an intact N-terminal tail, linker, and C-terminal region. These mutants contain either point mutations or deletions within the zinc finger domains identical to those which caused premature reverse

transcription *in vivo*. The point mutants contain a cysteine to histidine mutation which results in the conversion from a CCHC-type zinc finger to a CCCC zinc finger. Deletion mutants are lacking residues from the zinc fingers from the first cysteine to the last within the CCHC zinc-binding motif.

Binding to Structured and Unstructured Oligonucleotides Measured by Fluorescence Anisotropy

To investigate how mutating critical zinc binding residues within the zinc fingers or deleting the zinc fingers of the nucleocapsid region of Gag affects interactions with NAs, fluorescence anisotropy was used to monitor binding to both structured and unstructured fluorescently labeled RNA constructs. Gag's ability to bind specifically and non-specifically to NAs is crucial to its functions throughout the viral life-cycle. Binding to the highly structured ψ-packaging signal near the 5' end of the genome allows Gag to specifically select and package the viral genome from amongst a vast pool of cellular RNAs [4,19]. Non-specific binding is critical to Gag performing its other functions as a NA chaperone, most notably promoting the annealing of the primer $tRNA^{Lys3}$ to the viral genome [20-21]. NMR structures of NC in solution in complex with SL2 and SL3 from the packaging signal show that residues within the zinc fingers make specific contacts to exposed bases at the end of the stem-loops while non-specific binding can be attributed to interactions between the flanking basic regions and the backbone phosphates [19,22].

Non-specific binding was measured using GA-22. The FA results indicate that binding to unstructured oligonucleotides is largely unaffected, and in some cases even improved by mutation or deletion of the zinc fingers, as expected since the highly cationic N-terminus and linker remained intact. All mutants containing histidine to cysteine mutations and ΔZF2

exhibited affinity indistinguishable from that of wild-type while ΔZF1+2 and Znless Gag were slightly better at binding to GA-22 than Gag ΔP6. Surprisingly, ΔZF1 showed a dramatic improvement in non-specific binding, with an apparent K_D that is almost 9-fold lower than for wild-type and 5-fold lower than any of the other mutants.

Table 2. Binding Affinities of Gag Variants to Structured and Unstructured NAs

Protein Variant	K_d^{SL3} (nM)	K_d^{GA-22} (nM)
WT Gag ΔP6	80±17	281±35
H400C	68±11	262±72
H421C	96±15	233±125
H400,421C	300±7	225±103
ΔZF1	73±16	34±5
ΔZF2	266±58	254±35
ΔZF1+2	284±14	148±6
Znless Gag	611±214	164±48

Specific interactions between Gag or mutant Gag and NAs were examined by measuring binding to a 20-mer mimic of stem-loop 3 from the ψ-packaging signal. It was observed that Gag H400C and Gag H421C, but not Gag H400,421C, bound SL3 as well as wild-type Gag, in good standing with previous studies which have shown that these mutants package near wild-type levels of genomic RNA *in vivo*. In contrast, Gag ΔZF2, Gag ΔZF1+2, and EDTA-treated Znless Gag bound SL3 with reduce affinities, near those of non-specific levels. In the case of ΔZF1+2 and Znless Gag, the affinity for GA-22 was actually higher than for SL3, which would seem to indicate that these mutants have a stronger preference for single-stranded RNA than for double-stranded or hairpin regions. Unexpectedly, ΔZF1 showed similar affinity for the SL3 construct as WT Gag ΔP6. This is in contrast to numerous studies, which have shown the N-terminal zinc

finger to be more important for NC's chaperone activity and the finding that this same mutant packaged less than 5% genomic RNA compared with wild-type Gag [11].

Aggregation Ability is Weakly Dependent on Presence of Intact Zinc Fingers

A key aspect of NC or Gag's ability to act as a NA chaperone is the ability to form large, mobile aggregates of protein and NAs and thus facilitate duplex nucleation in the annealing of complementary sequences [6]. Aggregation activity provides a majority of the rate enhancement observed when compared with the uncatalyzed reaction [15]. This ability has mainly been attributed to the basic residues within the unstructured N-terminal tail of NC, as this positively charged region is able to neutralize electrostatic repulsion between annealing strands, as well as bind multiple strands at once, thus greatly increasing the likelihood of complementary strands interacting [6,15,23]. A sedimentation assay was used to assess the contribution of the zinc fingers within Gag to its aggregation ability. A plot of aggregated $tRNA^{Lys3}$ and ShortPBS RNA as a function of protein concentration was constructed and used to determine $K_{1/2,Agg}$, the protein concentration at which half-maximal aggregation was observed.

The ability to aggregate NAs was only weakly dependent on zinc finger structure. Over a half-hour time course, all proteins were able to achieve aggregation of nearly 100% at the concentrations tested. Mutating histidine to cysteine in only a single zinc finger or deleting a single zinc finger reduced aggregation ability by only 1.5- to 2-fold (Figure 4). Disruption of both zinc fingers by mutation, deletion, or EDTA treatment still had only a mild negative effect, raising the $K_{1/2,Agg}$ by only 3- to 4-fold (Figure 4). This moderate effect on aggregation was expected, since all of the mutant Gag constructs used in this study maintained intact N-terminal tails, the region of NC to which aggregation ability is mainly attributed [6].

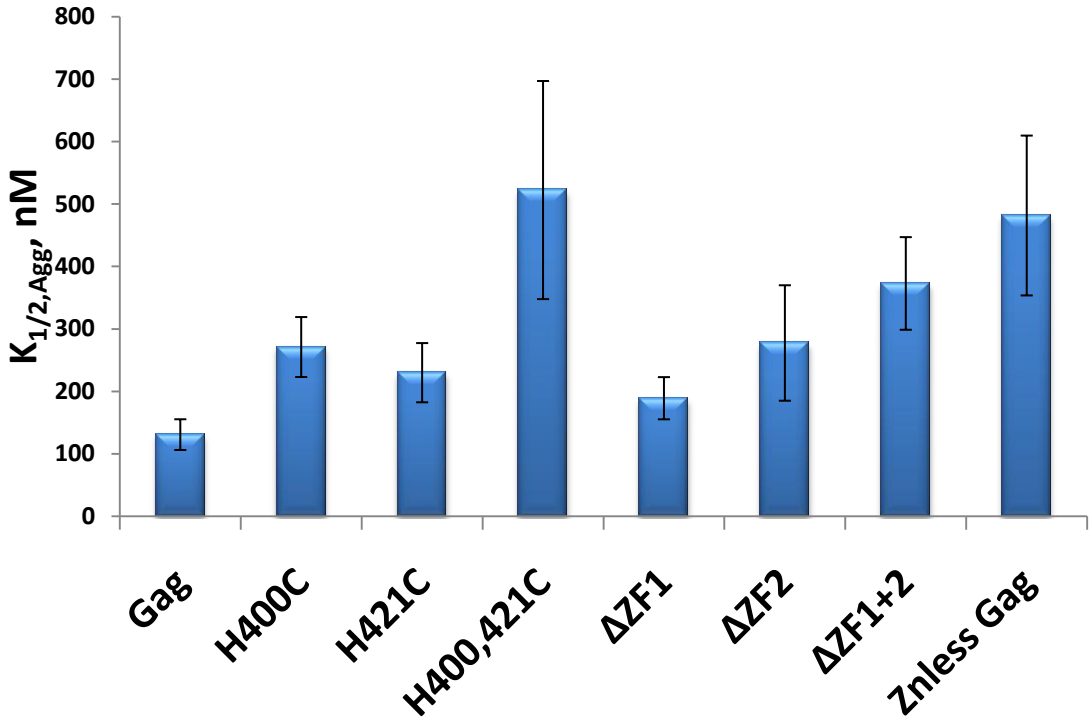

FIGURE 4. $K_{1/2,Agg}$ values determined for WT Gag and Gag zinc finger mutants.

Zinc Finger Mutants Maintain the Ability to Effectively Anneal the Primer tRNA^{Lys3} to the Genome

As an overall test of chaperone activity, the ability of Gag ΔP6 and the Gag zinc finger mutants to anneal *in vitro* transcribed tRNA^{Lys3} to the complementary Short PBS RNA was assessed. Steady-state annealing assays, in which a range of protein concentrations were tested over a period of 30 minutes, were used to assess how efficiently each protein catalyzed annealing. A plot of fraction annealed versus protein concentration was formulated for each protein and a $K_{1/2,Anneal}$ value, the protein concentration at which half-maximal annealing occurred was determined. For all mutants except H400,421C and Znless Gag, annealing was observed at similar protein concentrations, with the $K_{1/2,Anneal}$ values ranging from 128±44 nM for Gag ΔP6 to 237±26 nM for ΔZF1, with all other mutants falling in between (Figure 5).

Making an H to C mutation in a single zinc finger or deletion of a single or both zinc fingers led to less than a 2-fold increase in the concentration at which annealing was observed. H400,421C had a slightly higher loss in activity, with a $K_{1/2,Anneal}$ value of 360 nM, and Znless Gag had the most dramatic increase in $K_{1/2,Anneal}$, with a value of 655 nM (Figure 5).

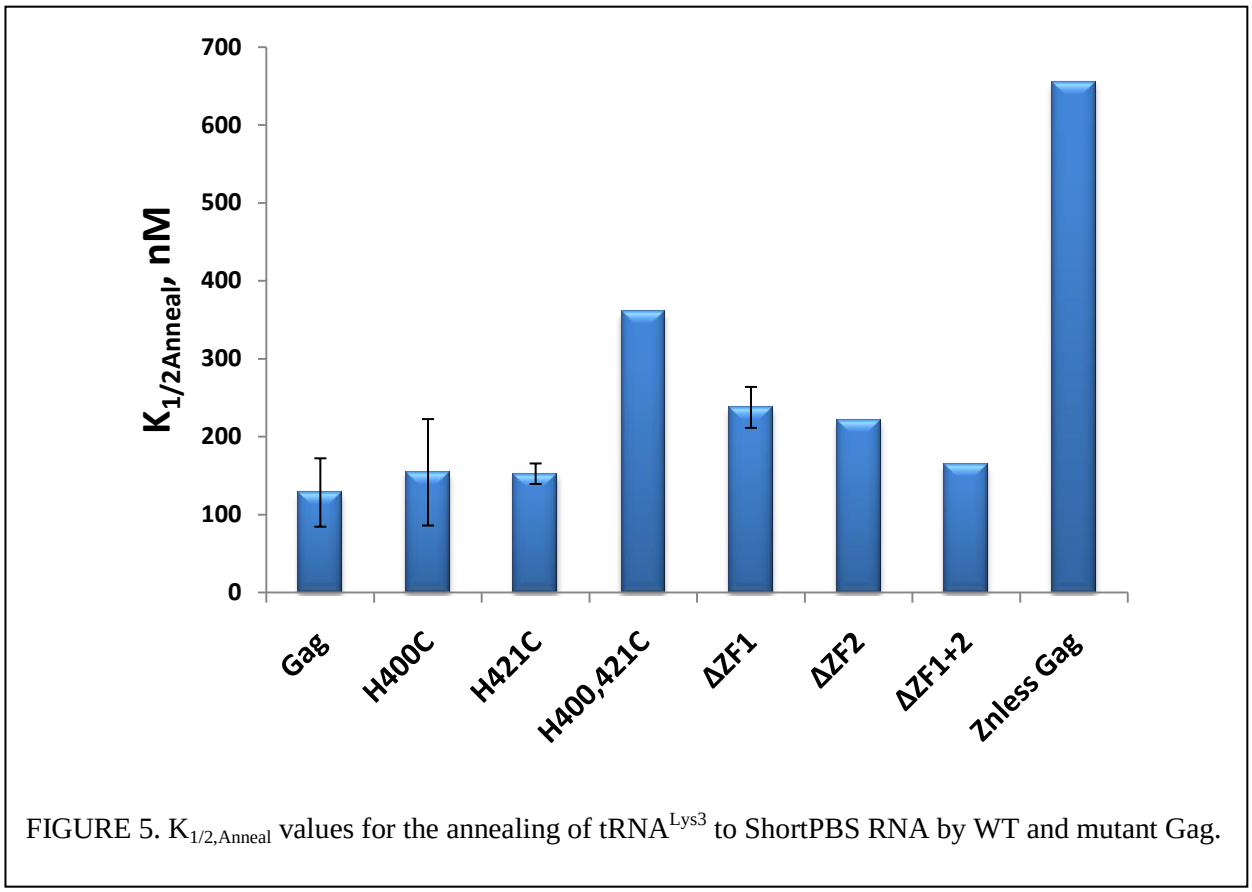

FIGURE 5. $K_{1/2,Anneal}$ values for the annealing of $tRNA^{Lys3}$ to ShortPBS RNA by WT and mutant Gag.

The results of the steady-state annealing assays were then used to determine the concentration of Gag or mutant Gag to use in monitoring the kinetics of the same reaction. In preliminary experiments, 400 nM protein was used for all constructs tested. However, no significant annealing was observed over the time-course tested (30 mins.) for H400,421C, ΔZF1+2, or Znless Gag, and an annealing rate k could not be reliably determined. Therefore, in subsequent experiments, 1 μM of each of these mutants was used. It is not presently clear why

ΔZF1+2 was not observed to effectively anneal $tRNA^{Lys3}$ to the Short PBS RNA at a concentration of 400 nM given the similar $K_{1/2,Anneal}$ value to Gag ΔP6 obtained from the steady-state assay. Under the conditions tested, mutants were capable of facilitating $tRNA^{Lys3}$ annealing (Figure 6). The highest rate of annealing was observed with the two mutants lacking the first zinc finger, ΔZF1 and ΔZF1+2, which had rates of 0.56±0.13/min. and 0.57±0.16/min., respectively

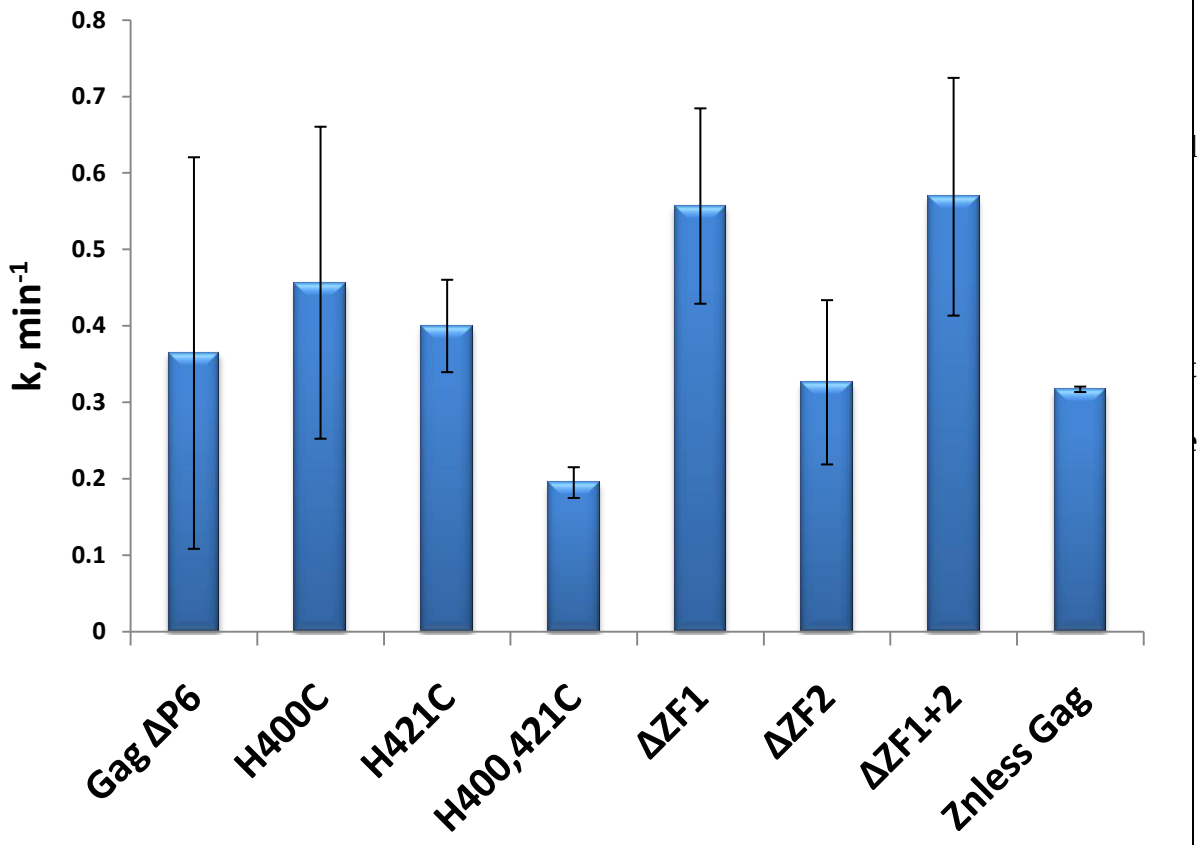

FIGURE 6. First-order rate constants for the annealing of $tRNA^{Lys3}$ to ShortPBS RNA by WT and mutant Gag variants.

genome and minus-strand transfer as well as markedly improves the processivity of reverse transcription by transiently eliminating elements of secondary structure which may cause pausing by reverse transcriptase.

Under the conditions tested here (See **Materials and Methods**), the free-standing NC domain showed no ability to disrupt the secondary structure of the TAR RNA. An additional study with the NC domain alone confirmed that even at concentrations up to 3.2 μM, NC was unable to destabilize the hairpin in 150 mM NaCl. However, at lower salt concentrations (20 or 50 mM), significant destabilization was observed by NC as previously reported [24-25]. Additionally, WT Gag ΔP6 showed minimal destabilization, as even at the highest concentration tested the relative free energy dropped from 1.8 kcal/mol to 1.4 kcal/mol, representing an

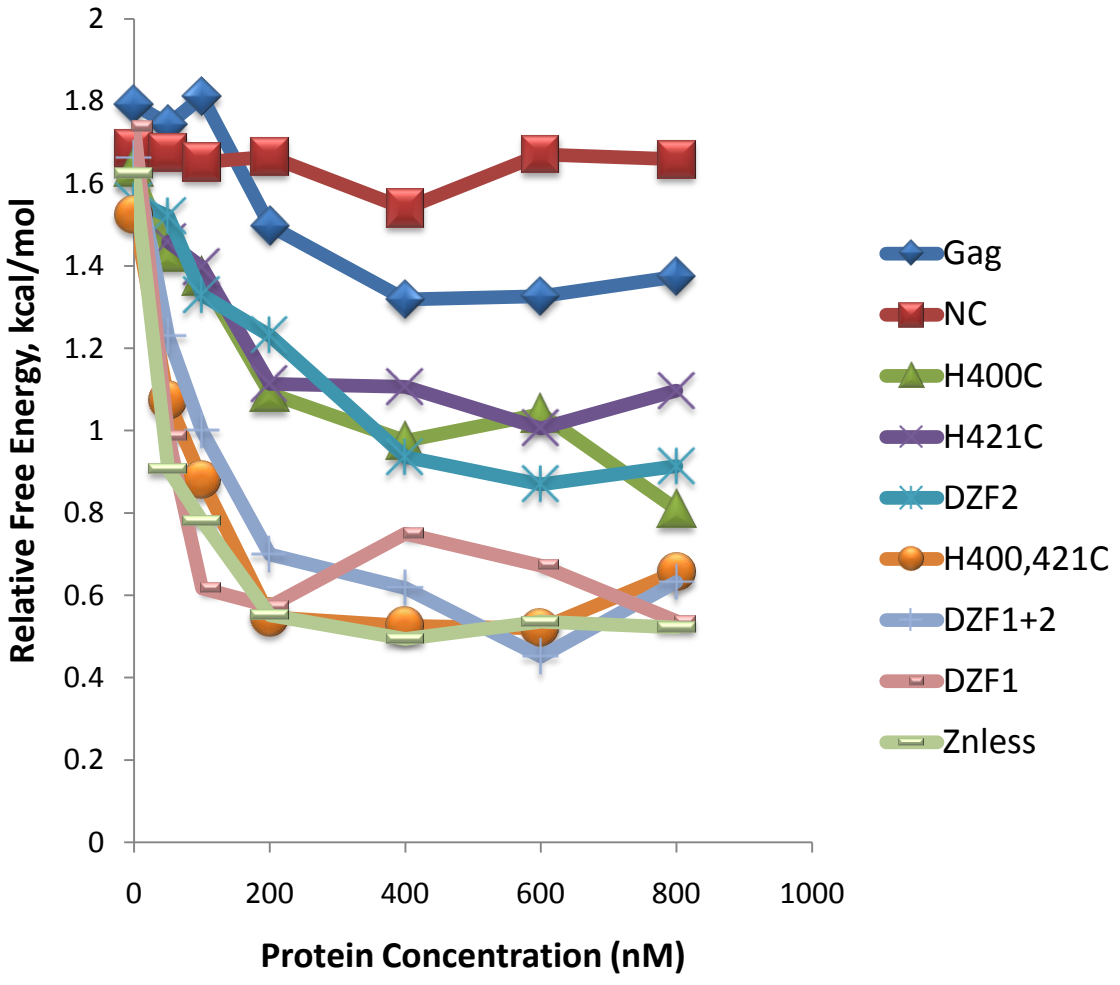

FIGURE 7. Relative free energy of the TAR DNA hairpin as a function of WT or mutant Gag variants or NC. Relative free energy calculated as ΔG=-RTln(% population in open state).

increase from ~6% in the fully open state in the absence of protein to ~11% in that state in the presence of saturating Gag. Surprisingly, the mutant Gag constructs tested showed a dramatic increase in their ability to destabilize NA secondary structure. At saturating mutant Gag concentrations, the presence of H400C, H421C, or ΔZF2 increased the percent of TAR molecules in the open state to 18-25% and in the presence of ΔZF1, ΔZF1+2, H400,421C, or Znless Gag the fully open state represented greater than 40% of the total population. Additionally, significant destabilization is observed at 50 nM ΔZF1, ΔZF1+2, H400,421C, and Znless Gag and 200 nM H400C, H421C, and ΔZF2, representing nucleotide:protein ratios of 128:1 and 32:1, respectively. Thus, mutation or removal of the zinc fingers of Gag appears to improve the destabilization capability.

Discussion

NC's NA chaperone activity is critical to the efficient replication of HIV-1 and it is thus of great importance to understand how NC interacts with NAs throughout the viral life cycle. This study investigated how mutations to the nucleocapsid domain within the context of Gag affect its chaperone activity. Binding to an SL3 mimic correlated well with data from *in vivo* studies investigating how these mutations affect genomic RNA (gRNA) packaging [11]. Those mutants which maintained wild-type affinities for SL3 exhibited moderate deficiencies in packaging, with H400C and H421C packaging reported to be 74% and 28% compared of wild-type, respectively [11]. In contrast, ΔZF1 and ΔZF2 packaged less than 10% compared with wild-type, while ΔZF1+2 packaged only 1% [12]. It is not clear why ΔZF1 showed reduced levels of packaging while still exhibiting wild-type affinity for SL3 in our *in vitro* assay. However, it is possible that the observed tighter binding to SL3 by ΔZF1 is explained by enhanced non-specific interactions, as evidenced by its extremely high affinity for GA-22. If this

were the case, ΔZF1 would likely package random mRNAs with similar frequency, thus resulting in lower gRNA per virion compared with wild-type. These binding studies further highlight Gag's role in packaging of the viral RNA and in general the *in vitro* results presented here correlate with those observed *in vivo* [26].

In accordance with previous studies indicating that the N-terminal basic residues are the main determinant for non-specific binding, aggregation of NAs [6], and annealing of $tRNA^{Lys3}$ [20], the Gag mutants studied here showed minimal defects in each of these activities. Even the 3- to 4-fold decrease in aggregation ability observed for the H400,421C, ΔZF1+2, and Znless mutants had essentially no effect on the rate of primer tRNA annealing to the genome. The annealing of $tRNA^{Lys3}$ to the PBS utilizes primarily the aggregation ability of NC or the NC domain of Gag [15], so it is possible that the stronger destabilization seen with these mutants compensated for the weaker aggregation. This contradicts what has been observed for zinc-less NC mutants, which exhibit weaker destabilization activity but are more proficient in promoting duplex nucleation [15]. Alternatively, the concentration of these proteins expressed *in vivo* may greatly surpass the threshold necessary for efficient aggregation to occur, making a 4-fold increase in the $K_{1/2}$ irrelevant, especially as the NC concentration in virions is estimated to be greater than 1 mM [27]. This indicates that the zinc-fingers within Gag are dispensable for the annealing of $tRNA^{Lys3}$ to the viral genome and that these mutants should be able to efficiently promote this *in vivo*.

So how might these mutant Gags be allowing for reverse transcription to occur prematurely in virus producing cells? One possibility is that the enhanced duplex destabilization ability observed for all of the mutants tested allows them to alter the conformation of the tRNA:gRNA complex so that reverse transcription occurs before budding. Studies in which

initiation of reverse transcription has been compared using templates extracted from wild-type or PR(-) virions have shown that $tRNA^{Lys3}$ annealed by Gag can not be extended by RT as readily as primer annealed by NC [28-29]. Additionally, these studies showed that transiently exposing Gag-annealed $tRNA^{Lys3}$ complexes led to initiation by RT comparable to that seen for PR(+) virions. Thus, it would seem that NC is able to promote the formation of a tRNA:gRNA complex that promotes initiation of reverse transcription while Gag lacks this ability.

A potential interaction that may be promoted by NC but not by Gag is the interaction between the primer activating signal (PAS), a block of eight nucleotides within the untranslated 5' (U5) leader sequence and the anti-PAS, the complementary sequence found within the TψC arm of $tRNA^{Lys3}$ [30]. This interaction has been shown to enhance reverse transcription initiation by disrupting the base pairing in the U5 leader stem. Mutations to the viral RNA, which cause the PAS to be constitutively single-stranded, lead to a significant increase in transcription initiation. It has also been demonstrated that a constitutively single-stranded PAS will allow for a substantial increase in reverse transcription initiation from primer:template complexes extracted from PR(-) virions, in which the tRNA is annealed by Gag, compared with PR(-) virions with wild-type genomic RNA [31]. Thus, destabilization of this leader sequence allows for reverse transcription to initiate. In this study, we showed that all of the mutants that allowed for reverse transcription to initiate prematurely within virus-producing cells had enhanced destabilization activity compared with wild-type Gag. Therefore, we hypothesize that the PAS sequence is being destabilized by mutant Gag and allowing reverse transcription to initiate. This seems likely given that the mutants tested were able to destabilize the TAR DNA hairpin at extremely low concentrations under physiological salt conditions, as low as 128:1 nt:protein, and as a result

would likely be able to promote this destabilization during assembly, when the pool of available dNTPs is sufficient for reverse transcription to proceed to completion.

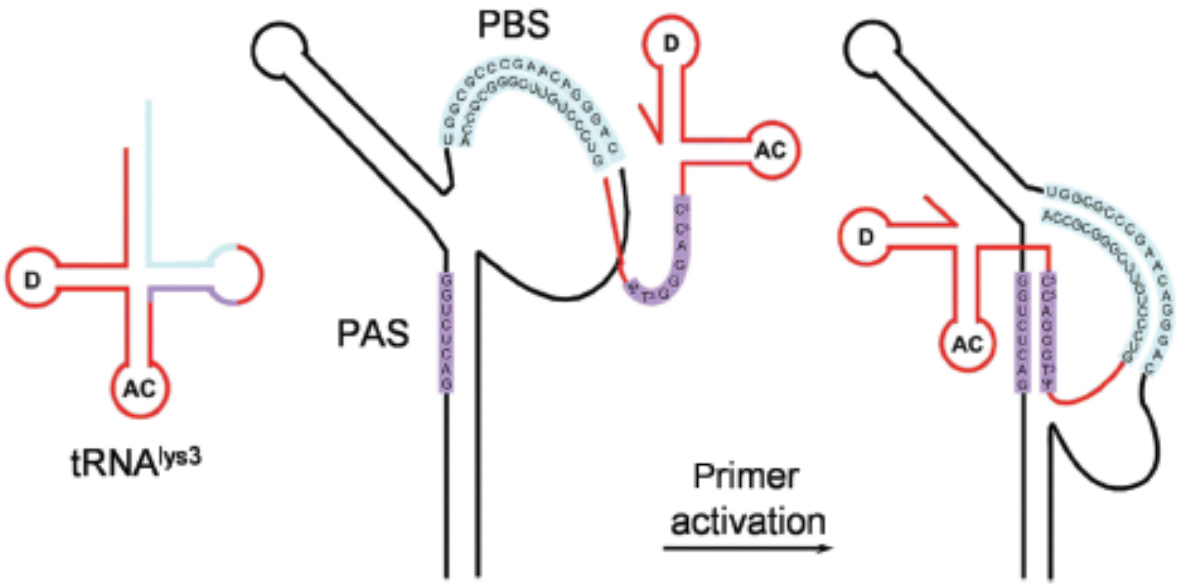

FIGURE 8. Schematic of the annealing of $tRNA^{Lys3}$ to the viral genome and subsequent rearrangement that takes place to form the PAS:anti-PAS interaction. It has been proposed that this interaction between the TψC arm of the primer tRNA and the viral genome is important for promoting the initiation of reverse transcription. Figure adapted from [3].

An alternative mechanism by which these Gag mutants promote reverse transcription is through the destabilization of the helix immediately 5' of the PBS (Fig. 9). This 8-bp helix may represent a significant obstacle that reverse transcriptase must overcome to continue synthesis of minus-strand strong-stop DNA []. Additionally, remodeling of this helix may be required for reverse transcriptase to switch from initiation to elongation mode. A model developed by Isel and colleagues of the $tRNA^{Lys3}$:gRNA:RT initiation complex showed that this helix interacted with the top and rear regions of the thumb and fingers subdomains of RT [32]. In contrast, previously proposed elongation models show the template strand lying in the cleft between these two subdomains [32]. Thus, we hypothesize the switch from initiation to elongation could be

promoted by a rearrangement of the template strand. Further support for this idea comes from Lanchy et al., who found that rate of incorporation of the fourth nucleotide by RT was extremely low compared with other nucleotides during initiation, likely due to the secondary structure imposed by this 8-bp helix [33]. Additionally, they demonstrated that the switch from initiation to elongation occurs between addition of the sixth and seventh nucleotides to the primer tRNA [33]. At this point, this helix should be sufficiently melted to no longer act as a barrier to RT and thus allow it to enter into a more processive mode.

This correlates well with recently proposed models of the secondary structure of the HIV-1 5' UTR and observations concerning natural endogenous reverse transcription [1,10]. The first study indicated that within the viral capsid, NC strongly destabilizes the first five nucleotides

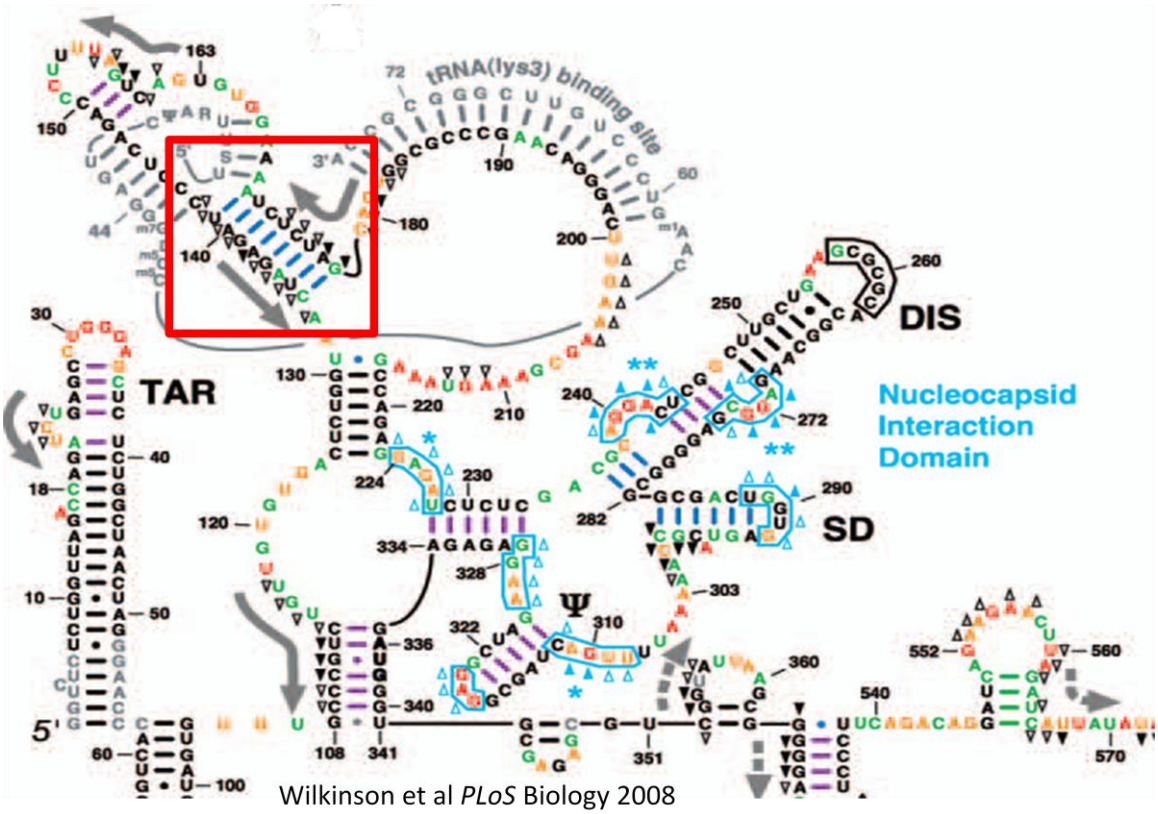

FIGURE 9. Proposed secondary structure of the 5' untranslated region of the HIV-1 genome. The red box highlights the 8-bp helix immediately ustream of the primer binding site (PBS) which may act to prevent premature reverse transcription. The grey arrows indicate the direction of reverse transcription, which begins from the $tRNA^{Lys3}$ primer shown in grey annealed to the genome. Adapted from [1].

acting as the template for RT and weakly destabilizes the sixth and seventh [1]. Furthermore, it was shown that removal of the lipid bi-layer from mature virions and addition of dNTPs to the medium allowed for extensive reverse transcription to occur, indicating that the conformation in which the genome exists within the mature capsid is competent for processive reverse transcription to occur [10]. This evidence suggests that RT relies on the high concentration of NC within the mature capsid to destabilize this helix in order to add the first six nucleotides, after which this secondary structure has been significantly destabilized and RT can switch to a conformation competent for elongation.

Conclusions and Future Direction

The results presented here show that the zinc fingers within the nucleocapsid domain of Gag are dispensable for Gag's functions as a nucleic acid chaperone outside of genome RNA packaging. Deletion of the zinc fingers led to lower affinity for our SL3 construct, reflecting reduced levels of genome packaging *in vivo*, while mutants containing only a single point mutation maintained wild-type affinity, also consistent with *in vivo* results. Mutations to the zinc fingers did not have any noticeable effect on non-specific binding to oligonucleotides, aggregation, or annealing of complementary sequences. Surprisingly, mutation of the zinc fingers led to an increase in the ability to destabilize nucleic acid secondary structure at very low Gag concentrations. Given the significant increase in destabilization activity at low protein concentrations observed for the Gag mutants that promote premature reverse transcription, it seems plausible that these mutants facilitate rearrangement of the 5' UTR of the viral genome at some point during assembly and thus lack the ability to prevent premature initiation of reverse transcription. However, additional studies need to be carried out before a clear picture of Gag's role in reverse transcription can be obtained. One unresolved question is how these mutants are

able to effectively destabilize nucleic acid secondary structure despite lacking competent zinc fingers, which numerous studies [6,24-25,34]] have shown are critical for this activity within the context of free-standing NC. Single-molecule DNA stretching should be able to determine the on/off kinetics of DNA binding and elucidate the mechanism by which these mutants achieve this function [8]. Additional studies will also be necessary to directly probe the interactions between Gag and the PAS and the 8-bp helix directly 5' of the PBS and how these interactions influence reverse transcription. First, it will need to be determined whether or not reverse transcription can occur *in vitro* in the presence of these Gag mutants. Preliminary results from our lab suggest that Gag blocks RT from extending $tRNA^{Lys3}$ annealed to the ShortPBS RNA, while NC does not. If it is determined that these mutants do not block extension by RT, further experiments to clarify Gag's role in initiation of reverse transcription will need to be carried out. Disruption of the 8-bp helix immediately downstream via mutation and the effects of these mutations on reverse transcription should indicate this structure's role in the viral life cycle. Previous studies found that disruption of this helix eliminated pausing by RT but also reduced the rate of transcription, suggesting moderate, but not total, destabilization of this structure is necessary for efficient reverse transcription [35]. Direct probing of Gag and mutant Gag interactions with this helix to establish whether they significantly destabilize it will also shed light on the role Gag is playing in regulation of reverse transcription. Finally, assessing Gag's role in facilitating the rearrangement necessary to establish the PAS:anti-PAS interaction and the influence this has on reverse transcription should further highlight Gag's role in influencing the timing of reverse transcription.

Acknowledgements

I would like to thank my adviser Dr. Karin Musier-Forsyth for extending me the

opportunity to become involved with undergraduate research and for her guidance throughout these past two years. My experience working in the lab along with her support and guidance has been invaluable to my education and development as a scientist while at Ohio State and prepared me to handle any challenges I may face after graduation.

I would also like to thank all the members of the Musier-Forsyth lab for their helpful advice, especially Chris Jones. His enthusiasm for teaching and endless encouragement and support were integral to my time at Ohio State and the passion for research and discovery he instilled in me is something I will carry throughout my life.

Finally, I would like to thank my parents for opening up so many opportunities throughout my life and to my brother and sister for their support throughout my college career.

Works Cited

[1] Wilkinson, K.A., Gorelick, R.J., Vasa, S.M., Guex, N., Rein, A., Mathews, D.H., Giddings, M.C. and Weeks, K.M. (2008). High-throughput SHAPE analysis reveals structures in HIV-1 genomic RNA strongly conserved across distinct biological states. PLoS Biol 6, e96.

[2] Ganser-Pornillos, B.K., Yeager, M. and Sundquist, W.I. (2008). The structural biology of HIV assembly. Curr Opin Struct Biol 18, 203-17.

[3] Ooms, M., Cupac, D., Abbink, T.E., Huthoff, H. and Berkhout, B. (2007). The availability of the primer activation signal (PAS) affects the efficiency of HIV-1 reverse transcription initiation. Nucleic Acids Res 35, 1649-59.

[4] Coffin, J.M., Hughes, S.H. and Varmus, H. (1997) Retroviruses, Cold Spring Harbor Laboratory Press. Plainview, N.Y.

[5] Rein, A., Henderson, L.E. and Levin, J.G. (1998). Nucleic-acid-chaperone activity of retroviral nucleocapsid proteins: significance for viral replication. Trends Biochem Sci 23, 297-301.

[6] Levin, J.G., Guo, J., Rouzina, I. and Musier-Forsyth, K. (2005). Nucleic acid chaperone activity of HIV-1 nucleocapsid protein: critical role in reverse transcription and molecular mechanism. Prog Nucleic Acid Res Mol Biol 80, 217-86.

[7] Herschlag, D. (1995). RNA chaperones and the RNA folding problem. J Biol Chem 270, 20871-4.

[8] Cruceanu, M., Gorelick, R.J., Musier-Forsyth, K., Rouzina, I. and Williams, M.C. (2006). Rapid kinetics of protein-nucleic acid interaction is a major component of HIV-1 nucleocapsid protein's nucleic acid chaperone function. J Mol Biol 363, 867-77.

[9] Feng, Y.X., Campbell, S., Harvin, D., Ehresmann, B., Ehresmann, C. and Rein, A. (1999). The human immunodeficiency virus type 1 Gag polyprotein has nucleic acid chaperone activity: possible role in dimerization of genomic RNA and placement of tRNA on the primer binding site. J Virol 73, 4251-6.

[10] Thomas, J.A., Bosche, W.J., Shatzer, T.L., Johnson, D.G. and Gorelick, R.J. (2008). Mutations in human immunodeficiency virus type 1 nucleocapsid protein zinc fingers cause premature reverse transcription. J Virol 82, 9318-28.

[11] Thomas, J.A., Gagliardi, T.D., Alvord, W.G., Lubomirski, M., Bosche, W.J. and Gorelick, R.J. (2006). Human immunodeficiency virus type 1 nucleocapsid zinc-finger mutations cause defects in reverse transcription and integration. Virology 353, 41-51.

[12] Houzet, L., Morichaud, Z., Didierlaurent, L., Muriaux, D., Darlix, J.L. and Mougel, M. (2008). Nucleocapsid mutations turn HIV-1 into a DNA-containing virus. Nucleic Acids Res 36, 2311-9.

[13] Didierlaurent, L., Houzet, L., Morichaud, Z., Darlix, J.L. and Mougel, M. (2008). The conserved N-terminal basic residues and zinc-finger motifs of HIV-1 nucleocapsid restrict the viral cDNA synthesis during virus formation and maturation. Nucleic Acids Res 36, 4745-53.

[14] Gorelick, R.J. et al. (1999). Strict conservation of the retroviral nucleocapsid protein zinc finger is strongly influenced by its role in viral infection processes: characterization of HIV-1 particles containing mutant nucleocapsid zinc-coordinating sequences. Virology 256, 92-104.

[15] Hargittai, M.R., Gorelick, R.J., Rouzina, I. and Musier-Forsyth, K. (2004). Mechanistic insights into the kinetics of HIV-1 nucleocapsid protein-facilitated tRNA annealing to the primer binding site. J Mol Biol 337, 951-68.

[16] Datta, S.A. and Rein, A. (2009). Preparation of recombinant HIV-1 gag protein and assembly of virus-like particles in vitro. Methods Mol Biol 485, 197-208.

[17] Stewart-Maynard, K.M., Cruceanu, M., Wang, F., Vo, M.N., Gorelick, R.J., Williams, M.C., Rouzina, I. and Musier-Forsyth, K. (2008). Retroviral nucleocapsid proteins display nonequivalent levels of nucleic acid chaperone activity. J Virol 82, 10129-42.

[18] Qualley, D.F., Stewart-Maynard, K.M., Wang, F., Mitra, M., Gorelick, R.J., Rouzina, I., Williams, M.C. and Musier-Forsyth, K. (2009). C-terminal domain modulates the nucleic acid chaperone activity of human T-cell leukemia virus type 1 (HTLV-1) nucleocapsid protein (NC) via an electrostatic mechanism. J Biol Chem

[19] De Guzman, R.N., Wu, Z.R., Stalling, C.C., Pappalardo, L., Borer, P.N. and Summers, M.F. (1998). Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element. Science 279, 384-8.

[20] Hargittai, M.R., Mangla, A.T., Gorelick, R.J. and Musier-Forsyth, K. (2001). HIV-1 nucleocapsid protein zinc finger structures induce tRNALys,3 structural changes but are not critical for primer/template annealing. J Mol Biol 312, 985-97.

[21] Demene, H. et al. (1994). 1H NMR structure and biological studies of the His23-->Cys mutant nucleocapsid protein of HIV-1 indicate that the conformation of the first zinc finger is critical for virus infectivity. Biochemistry 33, 11707-16.

[22] Amarasinghe, G.K., De Guzman, R.N., Turner, R.B., Chancellor, K.J., Wu, Z.R. and Summers, M.F. (2000). NMR structure of the HIV-1 nucleocapsid protein bound to stem-loop SL2 of the psi-RNA packaging signal. Implications for genome recognition. J Mol Biol 301, 491-511.

[23] Fisher, R.J. et al. (2006). Complex interactions of HIV-1 nucleocapsid protein with oligonucleotides. Nucleic Acids Res 34, 472-84.

[24] Hong, M.K., Harbron, E.J., O'Connor, D.B., Guo, J., Barbara, P.F., Levin, J.G. and Musier-Forsyth, K. (2003). Nucleic acid conformational changes essential for HIV-1 nucleocapsid protein-mediated inhibition of self-priming in minus-strand transfer. J Mol Biol 325, 1-10.

[25] Cosa, G., Harbron, E.J., Zeng, Y., Liu, H.W., O'Connor, D.B., Eta-Hosokawa, C., Musier-Forsyth, K. and Barbara, P.F. (2004). Secondary structure and secondary structure dynamics of DNA hairpins complexed with HIV-1 NC protein. Biophys J 87, 2759-67.

[26] Urbaneja, M.A., Kane, B.P., Johnson, D.G., Gorelick, R.J., Henderson, L.E. and Casas-Finet, J.R. (1999). Binding properties of the human immunodeficiency virus type 1 nucleocapsid protein p7 to a model RNA: elucidation of the structural determinants for function. J Mol Biol 287, 59-75.

[27] Thomas, J.A. and Gorelick, R.J. (2008). Nucleocapsid protein function in early infection processes. Virus Res 134, 39-63.

[28] Guo, F., Saadatmand, J., Niu, M. and Kleiman, L. (2009). Roles of Gag and NCp7 in facilitating tRNALys3 Annealing to viral RNA in human immunodeficiency virus type 1. J Virol 83, 8099-107.

[29] Cen, S., Khorchid, A., Gabor, J., Rong, L., Wainberg, M.A. and Kleiman, L. (2000). Roles of Pr55(gag) and NCp7 in tRNALys3 genomic placement and the initiation step of reverse transcription in human immunodeficiency virus type 1. J Virol 74, 10796-800.

[30] Beerens, N., Groot, F. and Berkhout, B. (2001). Initiation of HIV-1 reverse transcription is regulated by a primer activation signal. J Biol Chem 276, 31247-56.

[31] Saadatmand, J., Niu, M., Kleiman, L. and Guo, F. (2009). The contribution of the primer activation signal to differences between Gag- and NCp7-facilitated tRNALys3 annealing in HIV-1. Virology 391, 334-41.

[32] Isel, C., Westhof, E., Massire, C., Le Grice, S.F., Ehresmann, B., Ehresmann, C. and Marquet, R. (1999). Structural basis for the specificity of the initiation of HIV-1 reverse transcription. EMBO J 18, 1038-48.

[33] Lanchy, J.M., Ehresmann, C., Le Grice, S.F., Ehresmann, B. and Marquet, R. (1996). Binding and kinetic properties of HIV-1 reverse transcriptase markedly differ during initiation and elongation of reverse transcription. EMBO J 15, 7178-87.

[34] Liu, H.W. et al. (2005). Single-molecule FRET studies of important intermediates in the nucleocapsid-protein-chaperoned minus-strand transfer step in HIV-1 reverse transcription. Biophys J 89, 3470-9.

[35] Rong, L., Liang, C., Hsu, M., Guo, X., Roques, B.P. and Wainberg, M.A. (2001). HIV-1 nucleocapsid protein and the secondary structure of the binary complex formed between tRNA(Lys.3) and viral RNA template play different roles during initiation of (-) strand DNA reverse transcription. J Biol Chem 276, 47725-32.